



Eyre, H., & Lawry, J. (2015). A Dempster–Shafer model of imprecise assertion strategies: Workshop on Weighted Logics for AI - 2013. *Journal of Applied Logic*, 13(4 Part 1), 458-479.  
<https://doi.org/10.1016/j.jal.2014.12.003>

Peer reviewed version

License (if available):  
CC BY-NC-ND

Link to published version (if available):  
[10.1016/j.jal.2014.12.003](https://doi.org/10.1016/j.jal.2014.12.003)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Elsevier at <http://www.sciencedirect.com/science/article/pii/S1570868314000883>. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# A Dempster-Shafer Model of Imprecise Assertion Strategies

Henrietta Eyre,  
Black Swan Data,  
11 York Road, Tower Building,  
London SE1 7NX, UK  
email: henrietta.eyre@blackswan.com

Jonathan Lawry,  
Department of Engineering Mathematics,  
University of Bristol,  
Bristol BS8 1UB, UK  
email: j.lawry@bris.ac.uk

## Abstract

A Dempster-Shafer theory based model of assertion is proposed for multi-agent communications so as to capture both epistemic and strategic uncertainty. Treating assertion as a choice problem, we argue that for complex multi-agent communication systems, individual agents will only tend to have sufficient information to allow them to formulate imprecise strategies for choosing between different possible true assertions. In a propositional logic setting, an imprecise assertion strategy is defined as a functional mapping between a valuation and a set of true sentences, where the latter is assumed to contain the optimal assertion given that particular state of the world. Uncertainty is then quantified in terms of probability distributions defined on the joint space of valuations and strategies, naturally leading to Dempster-Shafer belief and plausibility measures on sets of possible assertions. This model is extended so as to include imprecise valuations and to provide a meta-level treatment of weak and strong assertions. As a case study, we consider the application of our proposed assertion models to the problem of choosing between a number of different vague descriptions, in the context of both epistemic and supervaluationist approaches to vagueness.

**Keywords** Multi-agent communication, assertion strategies, uncertainty, vagueness

## 1 Introduction

Any satisfactory theory of natural language communication must give an account of assertion. That is, how and why do we choose the particular assertions which we make in a given context? Certainly, veracity cannot be the only factor since, for a particular state of the world, there are usually a huge, perhaps infinite, number of grammatically correct true sentences available as possible candidates. But what are the additional factors at play and how can they best be modelled? Dating back to Lewis' work on *coordination games*

[18], game theory provides one established approach to the assertion problem. More recent contributions in this area include communication games proposed by Parikh [20] which are based on Grice’s maximums of cooperation in language [11]. Typically, communication games involve two players, a speaker and a listener, each with a payoff function depending on the decisions and actions taken as a result of interactions between agents <sup>1</sup>. In many cases there are optimal strategies for playing the games, taking the form of Nash equilibria. However, communication games which attempt to model the rich and varied multi-agent interactions which make up natural language communications in general, are likely to be highly complex. Individual agents involved in the game will tend to have only imprecise and uncertain knowledge both of the state of the world and also of the exact nature of the game itself. This will then make the identification of optimal strategies, if they exist, difficult or even impossible. Furthermore, given the dynamic nature of language there is an inherent problem with identifying a fixed optimal strategy. Although a communication strategy may be optimal for a fixed point in time we would need a more general strategy to be constantly adapting in order for it to remain optimal within a dynamical system. Instead, we suggest that in such situations agents will consider applying robust imprecisely defined strategies which take account of the best available evidence, but which are then unlikely to identify single optimal assertions. In this paper we propose a Dempster-Shafer theory approach to modelling imprecise assertion strategies in the presence of uncertainty both about the state of the world and about the exact nature of the communication game being played. It is important to note the distinction between these uncertainties. Indeed, Parikh [21] argues that the assertability of an expression must depend on both an agent’s belief in the truth of a sentence as well as factors external to that belief. Factors such as how a sentence is likely to be interpreted, recognition of differences in beliefs and motivations across a population and the consequences of any misinterpretations, may all contribute to uncertainty about the nature of a communication game. Here we consider only a very simple type of communication game in which the speaker identifies a single, one-off assertion to make to the listener. For instance, we do not attempt to model an interactive dialogue between the speaker and, potentially multiple, listeners, in which the speaker’s choice of her next assertion would need to take into account both her previous assertions and any responses made to them by the listeners. However, while the proposed model is preliminary and developed only for propositional logic, it does provide some initial insight into how imprecise probabilities, such as Dempster-Shafer belief and plausibility, could be applied to the assertion problem. Furthermore, we believe that this elementary model could be subsequently developed so to take account of a more interactive dialogue game, although such an extension lies beyond the scope of this current paper.

The assertion problem is becoming of increasing practical importance in areas of artificial intelligence including natural language generation systems [24], and language evolution

---

<sup>1</sup>See Allot [2] for a nice overview of game theoretic models of communication.

in robotics. Natural language generation systems are software tools which receive a representation of the state of the world as input, usually in terms of certain key attributes, and then output natural language text in the form of a description, summary, diagnosis or forecast. Existing systems include those developed for automatic weather forecasting [10] and as medical diagnosis tools [23]. The assertion problem is also of great importance in language games. This is a paradigm introduced by Steels [27] in order to study the evolution of communication protocols between agents in a simulated or real environment. As with communication games, language games are played between two agents. One agent, acting as a speaker, will formulate a linguistic utterance to be asserted to the interacting agent (the listener) given the agent’s conceptual model, a goal and the constraints placed upon this goal (for example by the environment). The listener must then recognise the assertion given to her, interpret its meaning in relation to her conceptual model and update this model so as to satisfy any constraints implied by the assertion.

In both the types of AI system described above, the assertion problem tends to be formulated as a kind of decision problem. For instance, in natural language generation systems, rather than viewing the problem as one of translation from some formal representation of the state description into natural language text, it is instead thought of as a problem of choosing between a number of alternative viable texts. Similarly, robotic agents playing a form of Steels’ language game must choose between different available descriptions of the objects they encounter. Here we adopt essentially the same approach, in which agents attempt to identify an optimal assertion in the face of uncertainty and partial information. In fact, the formulation of such a choice problem within a Dempster-Shafer theory setting requires the following unique assertion assumption on the part of the communicating agents: *Agents assume that for any given state of the world there is a single optimal assertion*<sup>2</sup>. This is required since our approach relies on an epistemic interpretation of Dempster-Shafer theory according to which the associated belief and plausibility measures quantify uncertainty about what is the single true state of the world. Now we might perhaps question the strength of this assumption, especially since, for example, it is easy to envisage payoff functions which give equal maximum value to several possible assertions in a given context. However, from a practical perspective, and in view of the fact that an agent must eventually identify a single assertion, it is perhaps reasonable for them to believe in the existence of a coherent optimal strategy which would allow them make such a choice in a principled manner.

An outline of the paper is as follows. Section 2 gives a brief introduction to Dempster-Shafer theory outlining its relationship with probability and imprecise probability. Section 3 proposes a formal definition of imprecise strategies in a propositional logic setting. Sec-

---

<sup>2</sup>We should clarify that under a model of probabilistic uncertainty an agent may generate a distribution over possible assertions. This means that in practice an agent will have a probability distribution over possible candidates for the single optimal assertion, and under a stochastic assertion rule any sentence with non-zero probability could be selected.

tion 4 then introduces a Dempster-Shafer model of assertion by allowing for both epistemic and strategic uncertainty. In section 5 we show how this model can be applied to the problem of choosing between different vague statements when adopting an epistemic approach to vagueness. The basic DS-model proposed in section 3 is extended in section 6 to allow for imprecisely specified states of world in addition to imprecise assertion strategies. Finally, section 7 gives some conclusions and discussion.

## 2 Dempster-Shafer Theory

In this section we give a short introduction to some of the basic ideas from the Dempster-Shafer theory of evidence [5], [25]. In contrast to probability theory, Dempster-Shafer theory identifies a pair of dual measures referred to as the belief and the plausibility respectively. The underlying intuition is that belief measures the level of evidence which *directly supports* a given hypothesis, whilst plausibility measures the level of evidence which is *at least consistent with* the hypothesis. Both measures are characterised by an underlying mass function or basic probability assignment which quantifies the level of support given exactly to each particular piece of evidence. A more formal treatment is then as follows:

Let  $\mathcal{U}$  denote a finite domain of discourse comprised of all possible states of the world. A mass function is a function  $m : 2^{\mathcal{U}} \rightarrow [0, 1]$  such that  $m(\emptyset) = 0$  and  $\sum_{A \subseteq \mathcal{U}} m(A) = 1$ . For  $A \subseteq \mathcal{U}$  we can now define the level of belief and plausibility in the hypothesis that the true state of the world is contained in  $A$ , as follows:

**Definition 1.** *Belief and Plausibility*

*Given a mass function  $m$  as above then the belief measure  $Bel : 2^{\mathcal{U}} \rightarrow [0, 1]$  and the plausibility measure  $Pl : 2^{\mathcal{U}} \rightarrow [0, 1]$  characterised by  $m$  are defined as follows:  $\forall A \subseteq \mathcal{U}$ ;*

$$Bel(A) = \sum_{B \subseteq A} m(B) \text{ and } Pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$$

*Notice that if  $m$  is non-zero only on singleton sets  $B = \{x\}$  for  $x \in \mathcal{U}$ , then  $m$  defines a probability distribution on  $\mathcal{U}$  and  $Bel = Pl$  is a probability measure on  $2^{\mathcal{U}}$ .*

From definition 1 it is easy to see that  $Bel$  and  $Pl$  are dual measures in the sense that,  $\forall A \subseteq \mathcal{U}$ ,  $Pl(A^c) = 1 - Bel(A)$ , and also satisfy  $Bel(\mathcal{U}) = Pl(\mathcal{U}) = 1$  and  $Bel(\emptyset) = Pl(\emptyset) = 0$ . Furthermore,  $Bel$  is a super-additive and  $Pl$  is a sub-additive measure so that; For  $A_1, \dots, A_m \subseteq \mathcal{U}$ ;

$$\begin{aligned} Bel\left(\bigcup_{i=1}^m A_i\right) &\geq \sum_{\emptyset \neq S \subseteq \{1, \dots, m\}} (-1)^{|S|-1} Bel\left(\bigcap_{i \in S} A_i\right) \text{ and} \\ Pl\left(\bigcap_{i=1}^m A_i\right) &\leq \sum_{\emptyset \neq S \subseteq \{1, \dots, m\}} (-1)^{|S|-1} Pl\left(\bigcup_{i \in S} A_i\right) \end{aligned}$$

In addition,  $Bel$  and  $Pl$  can be interpreted as a special case of lower and upper (imprecise) probabilities. To see this, let  $\mathcal{M}$  denote the set of all probability measures defined on  $2^{\mathcal{U}}$ . Now given a belief measure  $Bel$  we define a *credal set* [17]  $\mathcal{C} \subseteq \mathcal{M}$  according to:

$$\mathcal{C} = \{P \in \mathcal{M} : \forall A \subseteq \mathcal{U}, P(A) \geq Bel(A)\}$$

In this case,  $Bel$  and  $Pl$  correspond to the lower and upper probability measures induced by  $\mathcal{C}$  so that:

$$\forall A \subseteq \mathcal{U}, Bel(A) = \inf\{P(A) : P \in \mathcal{C}\} \text{ and } Pl(A) = \sup\{P(A) : P \in \mathcal{C}\}$$

Now if, in addition to the mass function  $m$ , there is also a prior probability measure  $P$  on  $2^{\mathcal{U}}$  then this naturally identifies a single posterior probability measure in the credal set  $\mathcal{C}$ .

**Definition 2.** *Posterior Probability*

Let  $P$  be a prior probability measure on  $2^{\mathcal{U}}$  and let the available evidence be characterised by the mass function  $m$ . Then the posterior measure  $P_m \in \mathcal{C}$  is given by:  $\forall A \subseteq \mathcal{U}$ ;

$$P_m(A) = \sum_{B \subseteq \mathcal{U}} P(A|B)m(B)$$

This gives the associated posterior probability distribution on  $\mathcal{U}$  as:

$$\forall x \in \mathcal{U}, P_m(x) = P(x) \sum_{B:x \in B} \frac{m(B)}{P(B)}$$

In the particular case that  $P$  is the uniform prior then  $P_m$  is the so called pignistic distribution [26] given by:

$$\forall x \in \mathcal{U}, P_m(x) = \sum_{B:x \in B} \frac{m(B)}{|B|}$$

Note that in order for  $P_m$  to be defined it is required that the prior  $P(B) > 0$  for all subsets  $B$  for which  $m(B) > 0$ , and is undefined otherwise. In addition, for any such prior  $P$ ,  $P_m \in \mathcal{C}$  since  $\forall A \subseteq \mathcal{U}$ ,

$$\begin{aligned} P_m(A) &= \sum_{B \subseteq \mathcal{U}} P(A|B)m(B) = \sum_{B \subseteq A} P(A|B)m(B) + \sum_{B \not\subseteq A} P(A|B)m(B) \\ &= \sum_{B \subseteq A} m(B) + \sum_{B \not\subseteq A} P(A|B)m(B) \geq \sum_{B \subseteq A} m(B) = Bel(A) \end{aligned}$$

Definition 2 can perhaps best be understood from the perspective of the random set interpretation of Dempster-Shafer theory [19], [5]. This requires us to think of the mass function  $m$  as the probability distribution of a random set  $\mathcal{E}$  into  $2^{\mathcal{U}}$ , and where  $\mathcal{E}$  has an associated probability measure  $\Lambda$  on  $2^{2^{\mathcal{U}}}$ , so that  $\forall A \subseteq \mathcal{U}$  the following holds;

$$\Lambda(\mathcal{E} = A) = m(A), \Lambda(\mathcal{E} \subseteq A) = Bel(A) \text{ and } \Lambda(\mathcal{E} \cap A \neq \emptyset) = Pl(A)$$

Within this interpretation we can view the posterior distribution given in definition 2 as resulting from a generalized form of Bayesian updating in which the evidence is itself uncertain. More formally, given a prior probability measure  $P$  on  $2^{\mathcal{U}}$ , suppose we then obtain uncertain evidence in the form of the random set  $\mathcal{E}$ . Now a Bayesian approach would require us to evaluate the posterior distribution  $P(x|\mathcal{E})$ , but this cannot be determined precisely since the set-value of  $\mathcal{E}$  is uncertain. However, if instead we take the expected value of  $P(x|\mathcal{E})$  across all the possible instantiations of  $\mathcal{E}$  then we obtain the posterior distribution  $P_m(x)$  as given in definition 2. That is;

$$\forall x \in \mathcal{U}, E_m(P(x|\mathcal{E})) = \sum_{B \subseteq \mathcal{U}} P(x|B) \Lambda(\mathcal{E} = B) = \sum_{B \subseteq \mathcal{U}} P(x|B) m(B) = P_m(x)$$

Smets [26] has proposed the pignistic distribution, a special case of definition 2 in which the prior is uniform, as an important decision making tool within Dempster-Shafer theory. Although it should be noted that Smets rejected the random set interpretation, instead justifying the pignistic distribution in terms of more general indifference principles. Indeed for different interpretations of Dempster-Shafer theory other mappings between belief functions and probability distributions can be justified, including, for example, using Dempster's rule to combine the prior and the mass function. See Daniel [4] for an overview of alternative probabilistic transformations.

### 3 Assertion Strategies

In this section we propose a formal definition of an assertion strategy within the simple context of propositional logic. Within this framework we then consider what is meant by precise and imprecise strategies.

Let  $\mathcal{L}$  be a language of propositional logic with propositional variables  $\mathcal{P} = \{p_1, \dots, p_n\}$  and connectives  $\wedge, \vee$  and  $\neg$ . Let  $S\mathcal{L}$  denote the sentences of  $\mathcal{L}$  generated in the standard way by recursive application of the connectives to the propositional variables, and let  $L\mathcal{L}$  denote the literals of  $\mathcal{L}$ . Let  $\mathbb{V}$  denote the set of all valuations of  $\mathcal{L}$ . We further assume that the set of admissible assertions is restricted to a finite subset of sentences  $A\mathcal{L} \subseteq S\mathcal{L}$ . The intuition behind this assumption is as follows: In principle an agent could assert any sentence in  $S\mathcal{L}$  which they believe to be true. However, we argue that in practice only a finite subset of sentences are actually ever assertible. Indeed, such a restriction could even be based mainly on syntactic considerations. For example, following a Gricean [11] maxim of quantity, agents may be unwilling to assert any sentence for which there is a syntactically simpler equivalent sentence<sup>3</sup> which they could assert in its place. This would mean, for instance, that neither  $\neg\neg p_1$  nor  $(p_1 \wedge p_2) \vee (p_1 \wedge \neg p_2)$  would ever be asserted in place of simply  $p_1$ .

---

<sup>3</sup>i.e. perhaps a sentence with fewer connectives.

While veracity is insufficient to provide a full account of assertion we do assume that, following Gricean principles [11], agents do at least want to be truthful. Consequently, a valuation identifies a maximal subset of possible assertions corresponding to those sentences in  $\mathcal{AL}$  which it designates as being true.

**Definition 3.** *Maximal Assertion Set*

A valuation  $v \in \mathbb{V}$  defines a maximal set of admissible assertions corresponding to those sentences in  $\mathcal{AL}$  which are true under  $v$ .

$$\mathcal{AL}(v) = \{\theta \in \mathcal{AL} : v(\theta) = 1\}$$

**Definition 4.** *Assertion Strategy*

An assertion strategy for  $\mathcal{L}$  is a function  $s : \mathbb{V} \rightarrow 2^{\mathcal{AL}} - \{\emptyset\}$  such that  $\forall v \in \mathbb{V}$ ,  $s(v) \subseteq \mathcal{AL}(v)$ . Let  $\mathbb{S}$  denote the set of all strategies for  $\mathcal{L}$ . It is convenient to think of a strategy  $s$  as being defined in a piecewise manner of the form: For  $\mathbb{V} = \{v_1, \dots, v_{2^n}\}$  and  $\{s_1, \dots, s_{2^n}\} \subseteq \mathbb{S}$ ;

$$s(v) = s_i(v) \text{ if } v = v_i$$

Let  $s_0$  denote the vacuous strategy for which  $s_0(v) = \mathcal{AL}(v)$ ,  $\forall v \in \mathbb{V}$ .

**Definition 5.** *Precise and Imprecise Strategies*

$s \in \mathbb{S}$  is a precise strategy for  $\mathcal{L}$  if  $\forall v \in \mathbb{V}$ ,  $|s(v)| = 1$ . A strategy  $s$  is precise for a particular valuation  $v$  if  $|s(v)| = 1$ . Let  $\mathbb{PS}$  denote the set of precise strategies and let  $\mathbb{PS}(v)$  denote the set of strategies which are precise for  $v$ . Notice that  $\mathbb{PS} = \bigcap_v \mathbb{PS}(v)$ . Also notice that any precise strategy  $s \in \mathbb{PS}$  can be written as a piecewise function  $s = s_i : v_i$  for  $i = 1, \dots, 2^n$  where  $s_i \in \mathbb{PS}(v_i)$ . We refer to any strategy in  $\mathbb{S} - \mathbb{PS}$  as an imprecise strategy.

## 4 Uncertainty and Assertion

As outlined above, a natural way of formulating the assertion decision problem is to assume that agents are playing a form of communication game in which they receive different payoffs for different assertions given different states of the world. The unique assertion assumption is then simply the assumption that there is a unique precise strategy  $s \in \mathbb{PS}$  for which the assertion  $s(v)$  will provide the speaking agent with the maximal payoff given the state of the world  $v \in \mathbb{V}$ . In this context an agent's uncertainty takes two main forms:

- *Epistemic Uncertainty:* This is uncertainty about what is the true state of the world. For a propositional model, epistemic uncertainty results from insufficient knowledge or evidence to determine which are the true propositional variables i.e. to identify the true valuation.



- *Strategic Uncertainty:* This is uncertainty about what is the optimal assertion strategy. We can think of this as resulting from a lack of knowledge about the exact nature of the communication game being played. For example, a speaker agent may be uncertain about the payoff functions and/or strategies of the listening agents.

Now as mentioned above the unique assertion assumption requires that there is a single optimal precise strategy  $s \in \mathbb{PS}$ , however, in the light of strategic uncertainty agents may only have sufficient evidence to support certain combinations of imprecise strategies. Hence, we initially propose an integrated model of both types of uncertainty in the form of a probability distribution  $w$  on  $\mathbb{V} \times \mathbb{S}$ <sup>4</sup>. This approach will then be extended so as to incorporate imprecisely defined valuations in section 6. Taking such a probability distribution as representing an agent's knowledge then this naturally generates a measure  $\mu : \mathcal{SL} \rightarrow [0, 1]$ , where  $\mu(\theta)$  quantifies the agent's belief that the sentence  $\theta$  is true.

**Definition 6.** *Given a probability distribution  $w$  on  $\mathbb{V} \times \mathbb{S}$  then  $\forall \theta \in \mathcal{SL}$ ;*

$$\mu(\theta) = \sum_{v:v(\theta)=1} w(v) \text{ where } w(v) = \sum_{s \in \mathbb{S}} w(v, s)$$

As given in definition 6,  $\mu$  is a probability measure defined on  $\mathcal{SL}$  and hence satisfies the following properties (see Paris [22] for an exposition):

- If  $\models \theta$  then  $\mu(\theta) = 1$ .
- If  $\theta \equiv \varphi$  then  $\mu(\theta) = \mu(\varphi)$ .
- If  $\models \neg(\theta \wedge \varphi)$  then  $\mu(\theta \vee \varphi) = \mu(\theta) + \mu(\varphi)$ .

In contrast, the agent's beliefs concerning what is the optimal assertion to make, as inferred from distribution  $w$ , are most naturally quantified by Dempster-Shafer belief and plausibility measures as follows:

**Definition 7.** *A DS model of Assertion*

*Given a probability distribution  $w$  on  $\mathbb{V} \times \mathbb{S}$ , let  $m : 2^{\mathcal{AL}} \rightarrow [0, 1]$  be a mass function on  $2^{\mathcal{AL}}$  such that:*

$$\forall A \subseteq \mathcal{AL} \ m(A) = \sum_{(v,s):s(v)=A} w(v, s)$$

*Let the corresponding belief and plausibility measures be denoted by  $Bel$  and  $Pl$  respectively. In this case, for  $B \subseteq \mathcal{AL}$ ,  $Bel(B)$  and  $Pl(B)$  respectively quantify the agent's belief and plausibility that the optimal assertion is contained in  $B$ .*

---

<sup>4</sup>In the sequel we will abuse notation and use the same symbol to present both a probability distribution and the probability measure which it induces.

Notice that in keeping with the random set interpretation of DS theory we can view that pair  $(v, s)$  as defining a random set  $s(v)$  into  $\mathcal{AL}$ . Also, the unique assertion assumption allows us to view  $\mathcal{AL}$  as a set of exclusive and exhaustive elements as required in DS theory, since according to this assumption there is a single optimal assertion and by definition it must be in  $\mathcal{AL}$ .<sup>5</sup>

The following theorem explores the relationship between the Dempster-Shafer belief and plausibility measures quantifying the agent's uncertainty about what is the optimal sentence to assert and the probability measure  $\mu$  quantifying the agent's uncertainty about what sentences are true.

**Theorem 8.** *Given a probability distribution  $w$  on  $\mathbb{V} \times \mathbb{S}$ , then for  $A \subseteq \mathcal{AL}$ ;*

$$Bel(A) \geq \mu\left(\bigvee_{\theta \in A} \theta \wedge \bigwedge_{\theta \in \mathcal{AL}-A} \neg\theta\right) \text{ and } Pl(A) \leq \mu\left(\bigvee_{\theta \in A} \theta\right)$$

Furthermore, if  $\sum_v w(v, s_0) = 1$  then;

$$Bel(A) = \mu\left(\bigvee_{\theta \in A} \theta \wedge \bigwedge_{\theta \in \mathcal{AL}-A} \neg\theta\right) \text{ and } Pl(A) = \mu\left(\bigvee_{\theta \in A} \theta\right)$$

*Proof.* Consider the marginal distribution on  $\mathbb{V}$  given by:

$$w(v) = \sum_{s \in \mathbb{S}} w(v, s)$$

By definition 6 we have that,  $\forall \theta \in \mathcal{AL}$ ;

$$\mu(\theta) = \sum_{v \in \mathbb{V}} w(v)$$

Now for any  $A \subseteq \mathcal{AL}$ ;

$$\begin{aligned} Bel(A) &= \sum_{B \subseteq A} m(B) = \sum_s \sum_{v: s(v) \subseteq A} w(v, s) \geq \sum_s \sum_{v: \mathcal{AL}(v) \subseteq A} w(v, s) \\ &= \sum_{v: \mathcal{AL}(v) \subseteq A} \sum_s w(v, s) = \sum_{v: \mathcal{AL}(v) \subseteq A} w(v) = w(\{v : v(\bigvee_{\theta \in A} \theta) = 1, v(\bigvee_{\theta \in \mathcal{AL}-A} \theta) = 0\}) \\ &= w(\{v : v(\bigvee_{\theta \in A} \theta \wedge \bigwedge_{\theta \in \mathcal{AL}-A} \neg\theta) = 1\}) = \mu\left(\bigvee_{\theta \in A} \theta \wedge \bigwedge_{\theta \in \mathcal{AL}-A} \neg\theta\right) \end{aligned}$$

Also,

$$\begin{aligned} Pl(A) &= \sum_{B \cap A \neq \emptyset} m(B) = \sum_s \sum_{v: s(v) \cap A \neq \emptyset} w(v, s) \leq \sum_s \sum_{v: \mathcal{AL}(v) \cap A \neq \emptyset} w(v, s) \\ &= \sum_{v: \mathcal{AL}(v) \cap A \neq \emptyset} \sum_s w(v, s) = \sum_{v: \mathcal{AL}(v) \cap A \neq \emptyset} w(v) = w(\{v : \exists \theta \in A, v(\theta) = 1\}) \\ &= w(\{v : v(\bigvee_{\theta \in A} \theta) = 1\}) = \mu\left(\bigvee_{\theta \in A} \theta\right) \end{aligned}$$

---

<sup>5</sup>Note that this exclusivity in no way requires the elements of  $\mathcal{AL}$  to be logically exclusive. Instead, it only means that exactly one of them is the correct assertion to make.

In the case that  $\sum_v w(v, s_0) = 1$  then if  $w(w, s) > 0$  it follows that  $s(v) = \mathcal{AL}(v)$ . Hence,

$$\begin{aligned} \sum_s \sum_{v:s(v) \subseteq A} w(v, s) &= \sum_s \sum_{v:\mathcal{AL}(v) \subseteq A} w(v, s) \text{ and} \\ \sum_s \sum_{v:s(v) \cap A \neq \emptyset} w(v, s) &= \sum_s \sum_{v:\mathcal{AL}(v) \cap A \neq \emptyset} w(v, s) \end{aligned}$$

as required.  $\square$

**Example 9.** Let  $\mathcal{P} = \{p_1, p_2, p_3\}$  and let  $\mathcal{AL} = \{p_1, p_2, \neg p_1, \neg p_2, p_1 \wedge p_2, p_1 \wedge \neg p_2, \neg p_1 \wedge p_2, \neg p_1 \wedge \neg p_2\}$ . Suppose a politician needs to identify an assertion to make to a particular audience. The proposition  $p_3$  indicates whether or not the audience is generally receptive to the politicians views.  $p_1$  and  $p_2$  are distinct positive statements forecasting the state of the economy in five years time. Now consider the following strategies:

- $s_1$ :  $s_1(v) = \mathcal{AL}(v) - \{\theta : \exists \varphi \in \mathcal{AL}(v), \varphi \models \theta, \varphi \neq \theta\}$ .
- $s_2$ :  $s_2(v) = \mathcal{AL}(v) - \{\varphi : \exists \theta \in \mathcal{AL}(v), \varphi \models \theta, \varphi \neq \theta\}$

Intuitively then  $s_1$  is the strategy of picking the most specific statements, whilst  $s_2$  is the strategy of picking the most general statements. Also notice that  $s_1$  is a precise strategy whilst  $s_2$  is an imprecise strategy. Suppose that the politician's knowledge is as given by a probability distribution  $w$  on  $\mathbb{V} \times \mathbb{S}$  as summarised in the following table.

valuation	$p_1$	$p_2$	$p_3$	strategy	$w$
$v_1$	1	1	1	$s_1$	0.25
$v_2$	1	0	1	$s_1$	0.15
$v_3$	0	1	1	$s_1$	0.05
$v_4$	0	0	1	$s_2$	0.05
$v_5$	1	1	0	$s_1$	0.25
$v_6$	1	0	0	$s_2$	0.15
$v_7$	0	1	0	$s_2$	0.05
$v_8$	0	0	0	$s_2$	0.05

Now

$$\begin{aligned} \mathcal{AL}(v_1) &= \mathcal{AL}(v_5) = \{p_1, p_2, p_1 \wedge p_2\} \\ \mathcal{AL}(v_2) &= \mathcal{AL}(v_6) = \{p_1, \neg p_2, p_1 \wedge \neg p_2\} \\ \mathcal{AL}(v_3) &= \mathcal{AL}(v_7) = \{\neg p_1, p_2, \neg p_1 \wedge p_2\} \\ \mathcal{AL}(v_4) &= \mathcal{AL}(v_8) = \{\neg p_1, \neg p_2, \neg p_1 \wedge \neg p_2\} \end{aligned}$$

Hence,

$$\begin{aligned} s_1(v_1) &= \{p_1 \wedge p_2\}, \quad s_1(v_2) = \{p_1 \wedge \neg p_2\}, \quad s_1(v_3) = \{\neg p_1 \wedge p_2\} \\ s_2(v_4) &= \{\neg p_1, \neg p_2\}, \quad s_1(v_5) = \{p_1 \wedge p_2\}, \quad s_2(v_6) = \{p_1, \neg p_2\} \\ s_2(v_7) &= \{\neg p_1, p_2\}, \quad s_2(v_8) = \{\neg p_1, \neg p_2\} \end{aligned}$$

From this we have the following mass function:

$$m := \{p_1 \wedge p_2\} : 0.5, \{p_1 \wedge \neg p_2\} : 0.15, \{\neg p_1 \wedge p_2\} : 0.05, \\ \{\neg p_1, \neg p_2\} : 0.1, \{p_1, \neg p_2\} : 0.15, \{\neg p_1, p_2\} : 0.05$$

Now consider  $A = \{p_1, \neg p_2, p_1 \wedge \neg p_2\}$ .

$$\begin{aligned} Bel(\{p_1, \neg p_2, p_1 \wedge \neg p_2\}) &= m(\{p_1, \neg p_2\}) + m(\{p_1 \wedge \neg p_2\}) \\ &= 0.15 + 0.15 = 0.3 \end{aligned}$$

and

$$\begin{aligned} Pl(\{p_1, \neg p_2, p_1 \wedge \neg p_2\}) &= m(\{p_1, \neg p_2\}) + m(\{p_1 \wedge \neg p_2\}) + m(\{\neg p_1, \neg p_2\}) \\ &= 0.15 + 0.15 + 0.1 = 0.4 \end{aligned}$$

Also,

$$\begin{aligned} &\bigvee_{\theta \in A} \theta \wedge \bigwedge_{\theta \in A\mathcal{L}-A} \neg \theta \\ &= (p_1 \vee \neg p_2 \vee (p_1 \wedge \neg p_2)) \wedge (\neg p_2 \wedge \neg(\neg p_1) \wedge \neg(p_1 \wedge p_2) \wedge \neg(\neg p_1 \wedge p_2) \wedge \neg(\neg p_1 \wedge \neg p_2)) \\ &\equiv (p_1 \wedge \neg p_2) \end{aligned}$$

Hence,

$$\begin{aligned} \mu(\bigvee_{\theta \in A} \theta \wedge \bigwedge_{\theta \in A\mathcal{L}-A} \neg \theta) &= \mu(p_1 \wedge \neg p_2) = w(\{v : v(p_1 \wedge \neg p_2) = 1\}) \\ &= w(v_2) + w(v_6) = 0.15 + 0.15 = 0.3 \end{aligned}$$

In addition,

$$\bigvee_{\theta \in A} \theta = p_1 \vee \neg p_2 \vee (p_1 \wedge \neg p_2) \equiv p_1 \vee \neg p_2$$

Hence,

$$\mu(\bigvee_{\theta \in A} \theta) = \mu(p_1 \vee \neg p_2) = 1 - w(v_3) - w(v_7) = 1 - 0.05 - 0.05 = 0.9$$

The above DS model of assertion requires us to determine a joint probability distribution on  $\mathbb{V} \times \mathbb{S}$ . This is potentially problematic since even for a moderately sized language the cardinality of this space is too large for us to feasibly determine probability values for all (but one of) its elements. Underlying this difficulty is the fact that  $|\mathbb{V}| = 2^n$  i.e. that the number of valuations of  $\mathcal{L}$  is exponential in the number of propositional variables. Furthermore, although we might expect the number of strategies under consideration to be much small than  $2^n$ , this may also increase with the language size. In order to partly

circumvent this problem we might consider adopting any one of a number of techniques which have been proposed in the literature for determining a probability distribution on valuations when given only incomplete information about the probability of certain sentences of  $\mathcal{L}$  i.e. when given a set of constraints on  $\mu$ . For example, Paris [22] provides a detailed analysis of different possible selection algorithms in the case that the constraints on  $\mu$  are linear. As Paris [22] points out such algorithms initially require us to check if the given set of constraints are consistent. This is referred to as *probabilistic satisfiability* and is NP hard in general, although Hansen et al. [12] overviews a number of linear programming based algorithms which can be applied in practice. Alternatively, we might assume that the propositional variables are independent of each other and then construct the distribution on  $\mathbb{V}$  by taking products of  $\mu(p_i)$  or  $\mu(\neg p_i)$  for  $i = 1, \dots, n$ . Such methods can allow us to determine the distribution on  $\mathbb{V}$ , but it still remains to construct the full joint distribution on  $\mathbb{V} \times \mathbb{S}$ . Rewriting  $w(v, s)$  as  $w(v)w(s|v)$  we might expect the conditional distribution  $w(s|v)$  to take a relatively simple form. In particular, for a given valuation  $v$  it is perhaps reasonable to assume that  $w(s|v) > 0$  only for a small number of strategies  $s$ . In other words, for a particular state of the world only a relatively small number of assertion strategies would be considered relevant. If this assumption were to hold then it would simplify the elicitation of the conditional distribution  $w(s|v)$ , which when combined with  $w(v)$  would then allow us to estimate the joint distribution  $w(v, s)$  as above.

#### 4.1 Assertion Probabilities

In the particular case that the agent's knowledge allows her to formulate precise strategies for all possible states of the world then any resulting uncertainty takes the form of a probability distribution on  $\mathcal{AL}$ . More formally, suppose that  $w$  on  $\mathbb{V} \times \mathbb{S}$  is such that,  $w(v, s) > 0$  implies that  $s \in \mathbb{PS}(v)$ , then the resulting mass function as given in definition 7 only allocates non-zero mass to singleton subsets of  $\mathcal{AL}$ . Consequently, in this case  $Bel = Pl$  corresponding to a probability measure on  $2^{\mathcal{AL}}$  with an associated probability distribution on  $\mathcal{AL}$  given by:

$$\forall \theta \in \mathcal{AL}, P_m(\theta) = \sum_v \sum_{s: s(v)=\{\theta\}} w(v, s)$$

An alternative scenario in which a probability distribution on assertions is naturally generated from a more general  $w$  on  $\mathbb{V} \times \mathbb{S}$  is when the agent also has an underlying prior probability distribution on  $\mathcal{AL}$ . Given such a prior then the additional knowledge  $w$  can be used to generate a posterior distribution as given in the following definition.

**Definition 10.** *Posterior Probability on Assertions*

*Let  $w$  be a probability distribution on  $\mathbb{V} \times \mathbb{S}$  and let  $P$  be a prior probability distribution*

on  $A\mathcal{L}$ . Then  $\forall \theta \in A\mathcal{L}$  let;

$$P_m(\theta) = P(\theta) \sum_{B \subseteq A\mathcal{L}: \theta \in B} \frac{m(B)}{P(B)} = \sum_{v \in \mathbb{V}} \sum_{s \in \mathbb{S}} P(\theta|s(v))w(v, s)$$

For example, suppose that  $A\mathcal{L} = L\mathcal{L}$  corresponding to the literals of  $\mathcal{L}$  then an agent might have a degree of prior preference for positive (i.e. non-negated) statements over negative statements but is otherwise a priori indifferent. This can be formalised by a prior where  $P(p_i) = \frac{\alpha}{n}$  and  $P(\neg p_i) = \frac{1-\alpha}{n}$  for  $i = 1, \dots, n$  (where  $\alpha \in [0, 1]$ ). Now suppose the agent gains additional knowledge about the state of the world but has no information about the optimal assertion strategy. This then takes the form of a distribution  $w$  for which  $\sum_v w(v, s_0) = 1$  and the posterior distribution given in definition 10 has the following form:  $\forall p_i \in \mathcal{P}$ ;

$$P_m(p_i) = \alpha \sum_{A \subseteq \mathcal{P}: p_i \in A} \frac{m(A \cup \{\neg p_i : p_i \in A^c\})}{\alpha|A| + (1-\alpha)(n - |A|)} \text{ and}$$

$$P_m(\neg p_i) = (1 - \alpha) \sum_{A \subseteq \mathcal{P}: p_i \notin A} \frac{m(A \cup \{\neg p_i : p_i \in A^c\})}{\alpha|A| + (1-\alpha)(n - |A|)}$$

In the case that  $\alpha = \frac{1}{2}$  then we have the following expression:  $\forall l \in L\mathcal{L}, P_m(l) = \frac{\mu(l)}{n}$ . Simplified models of this kind have been applied in language game studies [7], [8] and also in computing with words [13].

Notice that definition 10 is a special case of definition 2 in which a posterior probability distribution on the universe of discourse is determined from an underlying prior together with evidence in the form of a mass function.

In the current context, one natural interpretation of the posterior distribution given in definition 10 is that the agent is using the prior in order to specify precise strategies as restrictions of the imprecise strategies allocated non-zero probability by  $w$ . More formally, for  $s, s' \in \mathbb{S}$ , we say that  $s'$  is a restriction of  $s$ , denoted  $s' \leq s$  if  $\forall v \in \mathbb{V}, s'(v) \subseteq s(v)$ . Then the following theorem shows that we can view definition 10 in terms of a process in which given  $w$ , the agent determines a new distribution  $w'$  on  $\mathbb{V} \times \mathbb{S}$ , in which the probability  $w(v, s)$  is re-distributed to  $w'(v, s')$ , where  $s' \leq s$  and  $s' \in \mathbb{PS}(v)$ . Furthermore, this redistribution is done in such a way so as to be proportionate to the prior  $P$  on  $A\mathcal{L}$ . Notice that the mass function generated by  $w'$  as in definition 7 is restricted to singleton sets and hence naturally determines a probability distribution on assertions as described above. This distribution is equal to that given by definition 10 for the same prior.

**Theorem 11.** *Given  $w$  on  $\mathbb{V} \times \mathbb{S}$  and a prior  $P$  on  $A\mathcal{L}$ , let  $w'$  on  $\mathbb{V} \times \mathbb{S}$  be defined as follows: If  $s \in \mathbb{PS}(v)$  then let  $w'(v, s) = w(v, s)$ . Otherwise  $\forall \theta \in s(v)$  and  $\forall s' \leq s$  such that  $s'(v) = \{\theta\}$  let*

$$w'(s', v) = \frac{P(\theta|s(v))w(s, v)}{|\{s' \leq s : s'(v) = \{\theta\}\}|}$$

In this case:

$$P_{m'}(\theta) = P_m(\theta)$$

where  $m$  and  $m'$  are the mass functions on  $\mathcal{AL}$  generated by  $w$  and  $w'$  respectively.

*Proof.*

$$\begin{aligned}
\forall \theta \in \mathcal{AL}, P_{m'}(\theta) &= \sum_{v \in \mathbb{V}} \sum_{s \in \mathbb{S}} P(\theta|s(v))w'(v, s) \\
&= \sum_{v \in \mathbb{V}} \sum_{s \in \mathbb{PS}(v)} P(\theta|s(v))w(v, s) + \sum_{v \in \mathbb{V}} \sum_{s \in \mathbb{S} - \mathbb{PS}(v)} \sum_{s' \leq s: s'(v) = \{\theta\}} \frac{P(\theta|s(v))}{|\{s' \leq s : s'(v) = \{\theta\}\}|} w(v, s) \\
&= \sum_{v \in \mathbb{V}} \sum_{s \in \mathbb{PS}(v)} P(\theta|s(v))w(v, s) \\
&\quad + \sum_{v \in \mathbb{V}} \sum_{s \in \mathbb{S} - \mathbb{PS}(v)} P(\theta|s(v))w(v, s) \sum_{s' \leq s: s'(v) = \{\theta\}} \frac{1}{|\{s' \leq s : s'(v) = \{\theta\}\}|} \\
&= \sum_{v \in \mathbb{V}} \sum_{s \in \mathbb{PS}(v)} P(\theta|s(v))w(v, s) + \sum_{v \in \mathbb{V}} \sum_{s \in \mathbb{S} - \mathbb{PS}(v)} P(\theta|s(v))w(v, s) \\
&= \sum_{v \in \mathbb{V}} \sum_{s \in \mathbb{S}} P(\theta|s(v))w(v, s) = P_m(\theta)
\end{aligned}$$

□

**Example 12.** Consider the language and mass function given in example 9. Now further suppose that a prior  $P$  is defined on  $\mathcal{AL}$  such that:

$$\begin{aligned}
P(p_1) &= P(p_2) = 0.2, \quad P(\neg p_1) = P(\neg p_2) = 0.06, \quad P(p_1 \wedge p_2) = 0.25, \\
P(p_1 \wedge \neg p_2) &= 0.1, \quad P(\neg p_1 \wedge p_2) = 0.1, \quad P(\neg p_1 \wedge \neg p_2) = 0.03
\end{aligned}$$

Applying definition 10 to condition on the mass function  $m$  derived in example 9, we obtain a posterior distribution  $P_m$  as follows: Trivially, we have;

$$P_m(p_1 \wedge p_2) = 0.5, \quad P_m(p_1 \wedge \neg p_2) = 0.15, \quad P_m(\neg p_1 \wedge p_2) = 0.05, \quad P_m(\neg p_1 \wedge \neg p_2) = 0$$

In addition,

$$\begin{aligned}
P_m(\neg p_1) &= P(\neg p_1) \sum_{B: \neg p_1 \in B} \frac{m(B)}{P(B)} = P(\neg p_1) \left( \frac{m(\{\neg p_1, \neg p_2\})}{P(\{\neg p_1, \neg p_2\})} + \frac{m(\{\neg p_1, p_2\})}{P(\{\neg p_1, p_2\})} \right) \\
&= 0.06 \left( \frac{0.1}{0.06 + 0.06} + \frac{0.05}{0.2 + 0.06} \right) = 0.0615385
\end{aligned}$$

Similarly, we also obtain;

$$P_m(\neg p_2) = 0.0846154, \quad P_m(p_1) = 0.1153846, \quad P_m(p_2) = 0.0384615$$

It is perhaps appropriate at this juncture to consider the epistemic status of prior distributions defined on the set of possible assertions. If we adopt the unique assertion assumption that in any context there is a single optimal assertion to make, then from a Bayesian perspective we can use a prior probability distribution in order to represent knowledge which the speaker might have about what this assertion should be, before she obtains any evidence about strategies or the current state of the world. Such prior knowledge could be derived from general experience of socio-linguistic conventions, and take the form of rules-of-thumb for deciding what to say in the absence of any specific contextual information. For example, as above the speaker may believe a priori that it tends to be better to make positive rather than negative assertions. This belief might be captured as a parametrised prior giving higher weight to positive than negative literals, but which is otherwise uniform, or as in example 12, it might discriminate more precisely between different combinations of positive and negative statements. As is usually the case in the Bayesian approach, such probability distributions are simply a formalism for encoding prior knowledge or preferences. Furthermore, in DS theory as well as in imprecise probability theory more broadly, the validity of a precisely defined prior distribution is questioned, especially with regard to its failure to distinguish adequately between ignorance and uncertainty. However, while we would tend to accept this view, from a practical perspective the use of priors in assertion modelling can provide a convenient mechanism by which to encode speaker preferences between different types of assertions. In addition, posterior distributions as in definition 10 also give us an effective decision making criterion for when a single assertion needs to be selected.

## 5 Vagueness and Assertion

In this section we explore how imprecise assertion strategies can be applied in order to choose between different vague statements or descriptions. Here we associate vagueness with *semantic uncertainty* [16], which is taken as referring to uncertainty about the correct interpretation of certain defining predicates. For example, suppose that the predicate *short* is interpreted as an interval of heights  $[0, \epsilon]$  for some threshold value  $\epsilon > 0$ . In this simple case, semantic uncertainty would take the form of uncertainty about what is the correct value of  $\epsilon$ . From this perspective we might consider treating  $\epsilon$  as a random variable with an associated density function  $f$ , where the latter represents the agent's knowledge about how *short* should be correctly interpreted. This naturally results in a probabilistic version of membership function in which the membership of a height  $h$  in *short* corresponds to the probability that  $\epsilon \geq h$ . That is:

$$\mu_{short}(h) = P(\epsilon \geq h) = \int_h^{\infty} f(\epsilon) d\epsilon$$



A more general formulation of this idea was given in Lawry and Tang [15] and is based on a prototype theory representation for the extension of a vague predicate. Consider the unary predicates  $Q_i : i = 1, \dots, n$  with extensions corresponding to regions of an underlying space  $\Omega$  as follows: Let  $d : \Omega^2 \rightarrow \mathbb{R}^+$  be a pseudo-distance metric on  $\Omega$ , and let  $a_i \in \Omega$  be the prototype for  $Q_i$ . Then, according to [15], for any element  $x \in \Omega$ ,  $Q_i(x)$  holds if and only if  $d(x, a_i) \leq \epsilon_i$  where  $\epsilon_i$  is a boundary threshold for  $Q_i$ . Semantic uncertainty can then be represented by a joint probability density  $f$  on the threshold random variables  $\vec{\epsilon} = (\epsilon_1, \dots, \epsilon_n)$ . Now given  $x_1, \dots, x_n \in \Omega$  let  $p_i$  denote the proposition  $Q_i(x_i)$  for  $i = 1, \dots, n$ , then any given allocation  $\vec{\epsilon}$  of threshold values determines a valuation on the corresponding propositional language as follows:

$$v_{\vec{\epsilon}}(p_i) = \begin{cases} 1 : d(x_i, a_i) \leq \epsilon_i & \text{for } i = 1, \dots, n \\ 0 : \text{otherwise} \end{cases}$$

Given a joint density  $f$  on  $\vec{\epsilon}$  then this naturally generates a probability distribution  $w$  on  $\mathbb{V}$  according to:

$$\forall v \in \mathbb{V}, w(v) = P(\vec{\epsilon} : v_{\vec{\epsilon}} = v) = \int_{\vec{\epsilon} : v_{\vec{\epsilon}} = v} f(\vec{\epsilon}) d\epsilon$$

If an agent also has sufficient knowledge about the optimal assertion strategy for these propositions so as to determine a probability distribution on the joint space  $\mathbb{V} \times \mathbb{S}$  with the above distribution as a marginal, then the Dempster-Shafer model outlined in section 4 can be applied. The following example shows how our model could potentially be used and gives a simple illustration of what an assertion strategy might look like in this context.

**Example 13.** Consider the predicate *short* with extension  $[0, \epsilon]$  where  $\epsilon$  is a random variable into  $\mathbb{R}^+$  and where  $\epsilon$  is distributed according to a triangular distribution with density function  $f$  given by:

$$f = \begin{cases} 0 : \epsilon < 150 \\ \frac{1}{400}\epsilon - \frac{3}{8} : 150 \leq \epsilon < 170 \\ \frac{19}{40} - \frac{1}{400}\epsilon : 170 \leq \epsilon < 190 \\ 0 : \epsilon \geq 190 \end{cases}$$

The membership function  $\mu_{\text{short}} : \mathbb{R}^+ \rightarrow [0, 1]$  is then defined such that, for a given  $h \geq 0$ ,  $\mu_{\text{short}}(h)$  corresponds to the probability that  $h \in [0, \epsilon]$  i.e. the probability that *short*( $h$ ) is true. For this example, we have that:

$$\mu_{\text{short}}(h) = P(\epsilon \geq h) = \int_h^\infty f(\epsilon) d\epsilon = \begin{cases} 1 : h < 150 \\ -\frac{1}{800}h^2 + \frac{3}{8}h - \frac{217}{8} : 150 \leq h < 170 \\ \frac{1}{800}h^2 - \frac{19}{40}h + \frac{361}{8} : 170 \leq h < 190 \\ 0 : h \geq 190 \end{cases}$$

Now for the sequence of heights (in cm)  $h_1 = 155$ ,  $h_2 = 165$ ,  $h_3 = 175$  and  $h_4 = 185$ , let the proposition  $p_i$  denote ‘a person with height  $h_i$  is short’, i.e.  $p_i = \text{short}(h_i)$ , for  $i = 1, \dots, 4$  (see figure 13). In this case there are five possible valuations as given in table 1. The associated probability distribution  $w$  on  $\mathbb{V}$ , also shown in table 1, is determined as follows: All of the possible valuations have the property that if  $v(p_j) = 1$  then  $v(p_i) = 1$  for all  $i \leq j$ . Hence,

$$w(v) = P(\epsilon \in [h_j, h_{j+1})) = \mu_{\text{short}}(h_j) - \mu_{\text{short}}(h_{j+1}) \text{ where } j = \max\{i : v(p_i) = 1\}$$

Assuming that  $A\mathcal{L} = L\mathcal{L} = \{p_1, p_2, p_3, p_4, \neg p_1, \neg p_2, \neg p_3, \neg p_4\}$ , we now consider two scenarios regarding an agent’s knowledge about the optimal assertion strategy. In one case we assume the agent is completely ignorant about what strategy to apply and hence always applies the vacuous strategy  $s_0$ . In the second case we assume that the agent adopts the strategy  $s$  defined by:

$$s(v) = \{p_j, \neg p_{j+1}\} \text{ where } j = \max\{i : v(p_i) = 1\}$$

The intuition behind this strategy is that  $p_j$  and  $\neg p_{j+1}$  are the most informative positive and negative assertions respectively. This is because if  $p_j$  is asserted then the listener(s) can infer that  $p_i$  holds for all  $i \leq j$  and similarly if  $\neg p_{j+1}$  is asserted then they can infer that  $\neg p_i$  holds for all  $i \geq j + 1$ . Let  $w_1$  and  $w_2$  be probability distributions on  $\mathbb{V} \times \mathbb{S}$  both with marginal distribution  $w$  on  $\mathbb{V}$  as defined above, but where for  $w_1$  strategy  $s_0$  is always applied and for  $w_2$  strategy  $s$  is always applied. That is:

$$\forall v \in \mathbb{V}, w_1(v, s_0) = w_2(v, s) = w(v)$$

These two probability distributions generate mass functions  $m_1$  and  $m_2$  given by:

$$m_1 := \{p_1, p_2, p_3, p_4\} : 0.03125, \{p_1, p_2, p_3, \neg p_4\} : 0.25, \{p_1, p_2, \neg p_3, \neg p_4\} : 0.4375, \\ \{p_1, \neg p_2, \neg p_3, \neg p_4\} : 0.25, \{\neg p_1, \neg p_2, \neg p_3, \neg p_4\} : 0.03125$$

and

$$m_2 := \{p_4\} : 0.03125, \{p_3, \neg p_4\} : 0.25, \{p_2, \neg p_3\} : 0.4375 \\ \{p_1, \neg p_2\} : 0.25, \{\neg p_1\} : 0.03125$$

Now suppose that the agent has a uniform prior probability distribution on  $A\mathcal{L}$  then we can determine posterior distributions according to definition 10 as given in table 2. Notice that for  $P_{m_1}$  the assertions with maximal probability are  $p_1$  and  $\neg p_4$  which refer to the heights  $h_1$  and  $h_4$  with maximum membership values in *short* and  $\neg$ *short* respectively. In contrast  $P_{m_2}$  gives maximum probability to  $p_2$  and  $\neg p_3$  which refer to heights  $h_2$  and  $h_3$  these being the closest to the borderline between *short* and  $\neg$ *short*<sup>6</sup>.

---

<sup>6</sup>Here we think of borderline cases of *short* and  $\neg$ *short* as being those heights  $h$  for which  $\mu_{\text{short}}(h) \approx \mu_{\neg \text{short}}(h)$ , 170cm being the typical case.

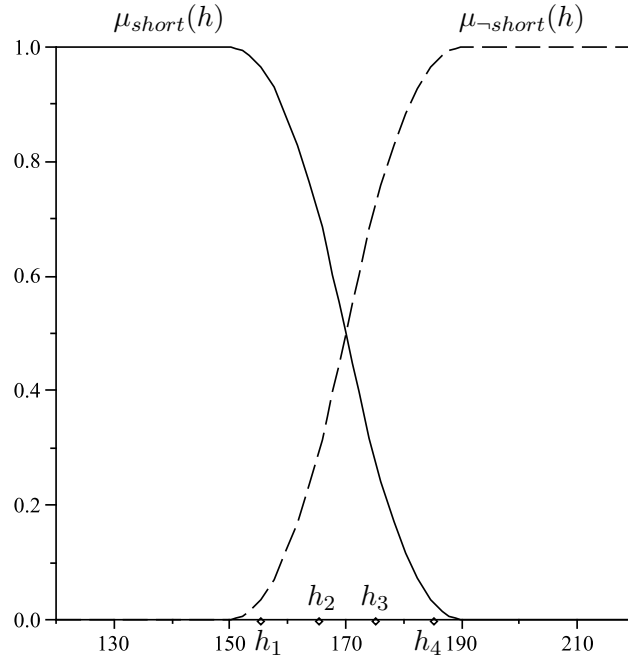


Figure 1: Plot showing the four heights  $h_1, \dots, h_4$  together with the membership functions for *short* and  $\neg$ *short*.

	$p_1$	$p_2$	$p_3$	$p_4$	$w(v)$	$s_0(v) = \mathcal{AL}(v)$	$s(v)$
$v_1$	1	1	1	1	$\mu_{short}(h_4) = 0.03125$	$\{p_1, p_2, p_3, p_4\}$	$\{p_4\}$
$v_2$	1	1	1	0	$\mu_{short}(h_3) - \mu_{short}(h_4) = 0.25$	$\{p_1, p_2, p_3, \neg p_4\}$	$\{p_3, \neg p_4\}$
$v_3$	1	1	0	0	$\mu_{short}(h_2) - \mu_{short}(h_3) = 0.4375$	$\{p_1, p_2, \neg p_3, \neg p_4\}$	$\{p_2, \neg p_3\}$
$v_4$	1	0	0	0	$\mu_{short}(h_1) - \mu_{short}(h_2) = 0.25$	$\{p_1, \neg p_2, \neg p_3, \neg p_4\}$	$\{p_1, \neg p_2\}$
$v_5$	0	0	0	0	$1 - \mu_{short}(h_1) = 0.03125$	$\{\neg p_1, \neg p_2, \neg p_3, \neg p_4\}$	$\{\neg p_1\}$

Table 1: Table showing the possible valuations generated by the membership function for the predicate *short*.

The probabilistic treatment of semantic uncertainty outlined in this section assumes an epistemic approach to vagueness. This brings us close to Williamson’s epistemic theory of vagueness [29] according to which any vague predicate has an objectively correct but uncertain boundary between it and its negation. However, we would argue the notion of semantic uncertainty requires only a somewhat weaker assumption about the epistemic nature of vagueness. In particular, it is sufficient to assume that agents, when faced with decision problems about assertions, find it useful as part of decision making strategy to simply assume that there is a correct crisp interpretation of the underlying predicates. In other words, when deciding what can be asserted agents behave as if the epistemic theory is correct. In earlier work we have referred to this strategic assumption across a population of agents as the epistemic stance [14]. In the next section we will consider an extension of the epistemic stance in which the underlying truth model is supervaluationist [9] rather

	$p_1$	$p_2$	$p_3$	$p_4$	$\neg p_1$	$\neg p_2$	$\neg p_3$	$\neg p_4$
$P_{m_1}$	0.24219	0.17969	0.07031	0.00781	0.00781	0.07031	0.17969	0.24219
$P_{m_2}$	0.125	0.21875	0.125	0.03125	0.03125	0.125	0.21875	0.125

Table 2: Table giving the posterior distributions on  $A\mathcal{L}$  generated by  $m_1$  and  $m_2$  assuming a uniform prior.

than classical.

## 6 Imprecise Valuations

In the previous sections, whilst we have allowed for imprecise strategies resulting from incomplete knowledge of the underlying communication game, we have assumed that the possible states of the world in the form of valuations on  $\mathcal{L}$  are precisely defined. We now weaken this assumption by allowing for imprecise valuations defined as subsets  $\Pi$  of  $\mathbb{V}$ . Imprecise valuations of this form can be completely characterised by associated upper valuations defined by:  $\bar{v} : S\mathcal{L} \rightarrow \{0, 1\}$  such that  $\forall \theta \in S\mathcal{L}$ ;

$$\bar{v}(\theta) = \max\{v(\theta) : v \in \Pi\}$$

In fact,  $\bar{v}$  is a boolean possibility measure on  $S\mathcal{L}$  [6]. We can also naturally define the following dual lower valuation corresponding to a Boolean necessity measure.

$$\underline{v}(\theta) = 1 - \bar{v}(\neg\theta) = \min\{v(\theta) : v \in \Pi\}$$

Let  $\mathbb{IV}$  be the set of all imprecise valuations of  $\mathcal{L}$ . We can then define a knowledge state as a probability distribution  $w$  on  $\mathbb{IV} \times \mathbb{S}$  where strategies in  $\mathbb{S}$  are extended from precise to imprecise valuations as follows: For  $s \in \mathbb{S}$ , and  $\Pi \in \mathbb{IV}$ ;

$$s(\Pi) = \bigcup_{v \in \Pi} s(v)$$

Also, we define the maximal assertion set of  $\Pi$  as;

$$A\mathcal{L}(\Pi) = s_0(\Pi) = \bigcup_{v \in \Pi} A\mathcal{L}(v)$$

Now a probability distribution on  $\mathbb{IV} \times \mathbb{S}$  naturally generates lower and upper measures  $\underline{\mu} : S\mathcal{L} \rightarrow [0, 1]$  and  $\bar{\mu} : S\mathcal{L} \rightarrow [0, 1]$  as follows:

**Definition 14.** *Given a probability distribution  $w$  on  $\mathbb{IV} \times \mathbb{S}$  then  $\forall \theta \in S\mathcal{L}$ ;*

$$\bar{\mu}(\theta) = \sum_{\Pi: \bar{v}(\theta)=1} \sum_s w(\Pi, s) \text{ and } \underline{\mu}(\theta) = \sum_{\Pi: \underline{v}(\theta)=1} \sum_s w(\Pi, s)$$

Here  $\underline{\mu}(\theta)$  and  $\overline{\mu}(\theta)$  quantify the agents subjective belief that  $\theta$  is *necessarily true* and *possibly true* respectively. Given this definition then in fact  $\underline{\mu}$  and  $\overline{\mu}$  correspond to Dempster-Shafer belief and plausibility measures on the sentences of  $\mathcal{L}$ , and consequently satisfy the following properties:  $\forall \theta, \varphi, \theta_1, \dots, \theta_m \in S\mathcal{L}$  (see [22] for an exposition);

- If  $\models \theta$  then  $\underline{\mu}(\theta) = \overline{\mu}(\theta) = 1$
- If  $\theta \equiv \varphi$  then  $\underline{\mu}(\theta) = \underline{\mu}(\varphi)$  and  $\overline{\mu}(\theta) = \overline{\mu}(\varphi)$ .
- $\overline{\mu}(\theta) = 1 - \underline{\mu}(\neg\theta)$ .
- $\underline{\mu}(\bigvee_{i=1}^m \theta_i) \geq \sum_{\emptyset \neq S \subseteq \{1, \dots, m\}} (-1)^{|S|-1} \underline{\mu}(\bigwedge_{i \in S} \theta_i)$  and  $\overline{\mu}(\bigwedge_{i=1}^m \theta_i) \leq \sum_{\emptyset \neq S \subseteq \{1, \dots, m\}} (-1)^{|S|-1} \overline{\mu}(\bigvee_{i \in S} \theta_i)$

On the other hand the agent's beliefs about what is the optimal assertion are now quantified by belief and plausibility measures  $Bel$  and  $Pl$  as characterised by the following mass function:

**Definition 15.** *Given a probability distribution  $w$  on  $\mathbb{IV} \times \mathbb{S}$  then  $\forall \theta \in S\mathcal{L}$ , let  $m : 2^{A\mathcal{L}} \rightarrow [0, 1]$  be such that:*

$$\forall A \subseteq A\mathcal{L}, \quad m(A) = \sum_s \sum_{\Pi: s(\Pi)=A} w(\Pi, s)$$

The following theorem explores the relationship between the lower and upper measures  $\underline{\mu}$  and  $\overline{\mu}$  on  $S\mathcal{L}$  quantifying the agents beliefs about which sentences are true and  $Bel$  and  $Pl$  measuring her beliefs about which is the optimal assertion to make.

**Theorem 16.** *Given a probability distribution  $w$  on  $\mathbb{IV} \times \mathbb{S}$  then for  $A \subseteq A\mathcal{L}$ ;*

$$Bel(A) \geq \underline{\mu}(\bigvee_{\theta \in A} \theta \wedge \bigwedge_{\theta \in A\mathcal{L}-A} \neg\theta) \text{ and } Pl(A) \leq \overline{\mu}(\bigvee_{\theta \in A} \theta)$$

Furthermore, if  $\sum_{\Pi} w(\Pi, s_0) = 1$  then;

$$Bel(A) = \underline{\mu}(\bigvee_{\theta \in A} \theta \wedge \bigwedge_{\theta \in A\mathcal{L}-A} \neg\theta) \text{ and } Pl(A) = \overline{\mu}(\bigvee_{\theta \in A} \theta)$$

*Proof.* Consider the marginal distribution on  $\mathbb{IV}$  given by;

$$w(\Pi) = \sum_s w(\Pi, s)$$

From this we have that;

$$\overline{\mu}(\theta) = \sum_{\Pi: \overline{v}(\theta)=1} w(\Pi) \text{ and } \underline{\mu}(\theta) = \sum_{\Pi: \underline{v}(\theta)=1} w(\Pi)$$

Now,

$$\begin{aligned}
Bel(A) &= \sum_{B \subseteq A} m(B) = \sum_s \sum_{\Pi: s(\Pi) \subseteq A} w(\Pi, s) \geq \sum_s \sum_{\Pi: A\mathcal{L}(\Pi) \subseteq A} w(\Pi, s) \\
&= \sum_{\Pi: A\mathcal{L}(\Pi) \subseteq A} \sum_s w(\Pi, s) = \sum_{\Pi: A\mathcal{L}(\Pi) \subseteq A} w(\Pi)
\end{aligned}$$

Now it holds that:

$$\begin{aligned}
A\mathcal{L}(\Pi) \subseteq A &\Leftrightarrow \bigcup_{v \in \Pi} A\mathcal{L}(v) \subseteq A \Leftrightarrow \forall v \in \Pi, A\mathcal{L}(v) \subseteq A \\
&\Leftrightarrow \forall v \in \Pi, v(\bigvee_{\theta \in A} \theta \wedge \bigwedge_{\theta \in A\mathcal{L}-A} \neg \theta) = 1 \Leftrightarrow \underline{v}(\bigvee_{\theta \in A} \theta \wedge \bigwedge_{\theta \in A\mathcal{L}-A} \neg \theta) = 1
\end{aligned}$$

Hence,

$$\sum_{\Pi: A\mathcal{L}(\Pi) \subseteq A} w(\Pi) = w(\{\Pi : \underline{v}(\bigvee_{\theta \in A} \theta \wedge \bigwedge_{\theta \in A\mathcal{L}-A} \neg \theta) = 1\}) = \underline{\mu}(\bigvee_{\theta \in A} \theta \wedge \bigwedge_{\theta \in A\mathcal{L}-A} \neg \theta)$$

Also,

$$\begin{aligned}
Pl(A) &= \sum_{B \cap A \neq \emptyset} m(B) = \sum_s \sum_{\Pi: s(\Pi) \cap A \neq \emptyset} w(\Pi, s) \leq \sum_s \sum_{\Pi: A\mathcal{L}(\Pi) \cap A \neq \emptyset} w(\Pi, s) \\
&= \sum_{\Pi: A\mathcal{L}(\Pi) \cap A \neq \emptyset} \sum_s w(\Pi, s) = \sum_{\Pi: A\mathcal{L}(\Pi) \cap A \neq \emptyset} w(\Pi)
\end{aligned}$$

Now

$$A\mathcal{L}(\Pi) \cap A \neq \emptyset \Leftrightarrow \exists \theta \in A, \exists v \in \Pi, v(\theta) = 1 \Leftrightarrow \exists v \in \Pi, v(\bigvee_{\theta \in A} \theta) = 1 \Leftrightarrow \overline{v}(\bigvee_{\theta \in A} \theta) = 1$$

Hence,

$$\sum_{\Pi: A\mathcal{L}(\Pi) \cap A \neq \emptyset} w(\Pi) = w(\{\Pi : \overline{v}(\bigvee_{\theta \in A} \theta) = 1\}) = \overline{\mu}(\bigvee_{\theta \in A} \theta)$$

Furthermore, in the case that  $\sum_{\Pi} w(\Pi, s_0) = 1$  we have that:

$$\begin{aligned}
\sum_s \sum_{\Pi: s(\Pi) \subseteq A} w(\Pi, s) &= \sum_s \sum_{\Pi: A\mathcal{L}(\Pi) \subseteq A} w(\Pi, s) \text{ and} \\
\sum_s \sum_{\Pi: s(\Pi) \cap A \neq \emptyset} w(\Pi, s) &= \sum_s \sum_{\Pi: A\mathcal{L}(\Pi) \cap A \neq \emptyset} w(\Pi, s)
\end{aligned}$$

as required.  $\square$

In the following example we show how the combination of imprecise valuations with imprecise assertion strategies can be relevant to the problem of choosing between different vague assertions. In particular, we now outline an extension of the model of vagueness

proposed in section 5 so as to incorporate imprecision as well as semantic uncertainty. Recall the unary predicates  $Q_i : i = 1, \dots, n$  defined as regions of a space  $\Omega$ , and with prototypes  $a_i \in \Omega$  for  $i = 1, \dots, n$ . Now instead of a single boundary threshold  $\epsilon_i$  for each  $Q_i$  it is supposed that there are lower and upper thresholds  $0 < \underline{\epsilon}_i \leq \bar{\epsilon}_i$  in each case. Since any tuple of precise boundary values  $\vec{\epsilon}$  naturally generates a precise valuation  $v_{\vec{\epsilon}}$  as defined in section 5, then a tuple of lower and upper bounds  $\vec{\rho} = (\underline{\epsilon}_1, \bar{\epsilon}_1, \dots, \underline{\epsilon}_n, \bar{\epsilon}_n)$  naturally generates an imprecise valuation of the form:

$$\Pi_{\vec{\rho}} = \{v_{\vec{\epsilon}} : \underline{\epsilon}_i \leq \epsilon_i \leq \bar{\epsilon}_i, i = 1, \dots, n\}$$

The associate lower and upper valuations  $\underline{v}_{\vec{\rho}}$  and  $\bar{v}_{\vec{\rho}}$  then satisfy the following properties:

$$\underline{v}_{\vec{\rho}}(p_i) = \begin{cases} 1 : d(x_i, a_i) \leq \underline{\epsilon}_i \\ 0 : \text{otherwise} \end{cases} \quad \text{and} \quad \bar{v}_{\vec{\rho}}(p_i) = \begin{cases} 1 : d(x_i, a_i) \leq \bar{\epsilon}_i \\ 0 : \text{otherwise} \end{cases} \quad \text{for } i = 1, \dots, n$$

This approach to modelling vagueness is a form of supervaluationism as proposed by Fine [9]. From this perspective the imprecise valuation  $\Pi_{\vec{\rho}}$  identifies a set of admissible precise (classical) valuations or precisifications. In other words, the inherent vagueness of the language means that there are a number of valid precise interpretations which can be appropriately applied. The formulation of supervaluationism for propositional languages in terms of lower and upper valuations is described in detail in [16].

Now suppose that in addition to imprecise valuations there is also semantic uncertainty. In the current context this might take the form of uncertainty about the lower and upper threshold values. For instance, treating the lower and upper thresholds as random variables and given a joint probability density function  $f$  on  $\vec{\rho}$  we can naturally generate a probability distribution on  $\mathbb{IV}$  given by:

$$w(\Pi) = P(\vec{\rho} : \Pi_{\vec{\rho}} = \Pi) = \int_{\vec{\rho} : \Pi_{\vec{\rho}} = \Pi} f(\vec{\rho}) d\rho$$

Furthermore, if  $f_i$  is the corresponding marginal distribution on  $(\underline{\epsilon}_i, \bar{\epsilon}_i)$  then we can determine lower and upper membership functions for the predicate  $Q_i$  according to:

$$\begin{aligned} \underline{\mu}_{Q_i}(x) &= P(\underline{\epsilon}_i \geq d(x, a_i)) = \int_{d(x, a_i)}^{\infty} \int_{\underline{\epsilon}_i}^{\infty} f_i(\underline{\epsilon}_i, \bar{\epsilon}_i) d\bar{\epsilon}_i d\underline{\epsilon}_i \text{ and} \\ \bar{\mu}_{Q_i}(x) &= P(\bar{\epsilon}_i \geq d(x, a_i)) = \int_{d(x, P_i)}^{\infty} \int_0^{\bar{\epsilon}_i} f_i(\underline{\epsilon}_i, \bar{\epsilon}_i) d\underline{\epsilon}_i d\bar{\epsilon}_i \end{aligned}$$

**Example 17.** Suppose that the predicate *short* is imprecisely defined in terms of lower and upper thresholds  $\underline{\epsilon} \leq \bar{\epsilon}$  on heights, so that any extension of *short* contained in  $\{[0, \epsilon] : \underline{\epsilon} \leq \epsilon \leq \bar{\epsilon}\}$  is admissible. Furthermore, suppose that there is uncertainty about the exact values of  $\underline{\epsilon}$  and  $\bar{\epsilon}$ , and that the agent's beliefs about these thresholds is represented by a joint probability density function  $f$  on  $(\underline{\epsilon}, \bar{\epsilon})$  satisfying:

$$\int_0^{\infty} \int_{\underline{\epsilon}}^{\infty} f(\underline{\epsilon}, \bar{\epsilon}) d\bar{\epsilon} d\underline{\epsilon} = 1$$

$\Pi$	$w(\Pi)$	$s(\Pi)$
$\Pi_1 = \{v_5, v_4, v_3\}$	$(1 - \underline{\mu}_s(h_1))(1 - \overline{\mu}_s(h_3)) = 0.015625$	$\{p_1, p_2, \neg p_1, \neg p_2, \neg p_3\}$
$\Pi_2 = \{v_5, v_4, v_3, v_2\}$	$(1 - \underline{\mu}_s(h_1))(\overline{\mu}_s(h_3) - \overline{\mu}_s(h_4)) = 0.09375$	$\{p_1, p_2, p_3, \neg p_1, \neg p_2, \neg p_3, \neg p_4\}$
$\Pi_3 = \{v_5, v_4, v_3, v_2, v_1\}$	$(1 - \underline{\mu}_s(h_1))\overline{\mu}_s(h_4) = 0.015625$	$\{p_1, p_2, p_3, p_4, \neg p_1, \neg p_2, \neg p_3, \neg p_4\}$
$\Pi_4 = \{v_4, v_3\}$	$(\underline{\mu}_s(h_1) - \underline{\mu}_s(h_2))(1 - \overline{\mu}_s(h_3)) = 0.09375$	$\{p_1, p_2, \neg p_2, \neg p_3\}$
$\Pi_5 = \{v_4, v_3, v_2\}$	$(\underline{\mu}_s(h_1) - \underline{\mu}_s(h_2))(\overline{\mu}_s(h_3) - \overline{\mu}_s(h_4)) = 0.5625$	$\{p_1, p_2, p_3, \neg p_2, \neg p_3, \neg p_4\}$
$\Pi_6 = \{v_4, v_3, v_2, v_1\}$	$(\underline{\mu}_s(h_1) - \underline{\mu}_s(h_2))\overline{\mu}_s(h_4) = 0.09375$	$\{p_1, p_2, p_3, p_4, \neg p_2, \neg p_3, \neg p_4\}$
$\Pi_7 = \{v_3\}$	$\underline{\mu}_s(h_2)(1 - \overline{\mu}_s(h_3)) = 0.015625$	$\{p_2, \neg p_3\}$
$\Pi_8 = \{v_3, v_2\}$	$\underline{\mu}_s(h_2)(\overline{\mu}_s(h_3) - \overline{\mu}_s(h_4)) = 0.09375$	$\{p_2, p_3, \neg p_3, \neg p_4\}$
$\Pi_9 = \{v_3, v_2, v_1\}$	$\underline{\mu}_s(h_2)\overline{\mu}_s(h_4) = 0.015625$	$\{p_2, p_3, p_4, \neg p_3, \neg p_4\}$

Table 3: Table showing the possible imprecise valuations generated by the lower and upper membership functions for the predicate *short*.

Based on  $f$  we can naturally define lower and upper membership functions of *short* so that for any height  $h$  (see figure 2);

$$\begin{aligned}\underline{\mu}_{short}(h) &= P(\underline{\epsilon} \geq h) = \int_h^\infty \int_{\underline{\epsilon}}^\infty f(\underline{\epsilon}, \overline{\epsilon}) \, d\overline{\epsilon} \, d\underline{\epsilon} \text{ and} \\ \overline{\mu}_{short}(h) &= P(\overline{\epsilon} \geq h) = \int_h^\infty \int_0^{\overline{\epsilon}} f(\underline{\epsilon}, \overline{\epsilon}) \, d\underline{\epsilon} \, d\overline{\epsilon}\end{aligned}$$

Now suppose that in this case the agent believes that  $\underline{\epsilon}$  and  $\overline{\epsilon}$  are independent variables both with triangular distributions centered around 160cm and 180cm respectively. More specifically;  $f(\underline{\epsilon}, \overline{\epsilon}) = f_1(\underline{\epsilon}) \times f_2(\overline{\epsilon})$  where

$$f_1(\underline{\epsilon}) = \begin{cases} \frac{\underline{\epsilon}-150}{100} : \underline{\epsilon} \in [150, 160] \\ \frac{160-\underline{\epsilon}}{100} : \underline{\epsilon} \in [160, 170] \\ 0 : \text{otherwise} \end{cases} \quad \text{and} \quad f_2(\overline{\epsilon}) = \begin{cases} \frac{\overline{\epsilon}-170}{100} : \overline{\epsilon} \in [170, 180] \\ \frac{190-\overline{\epsilon}}{100} : \overline{\epsilon} \in [180, 190] \\ 0 : \text{otherwise} \end{cases}$$

Now consider the propositions  $p_1, \dots, p_4$  generated by the four heights  $h_1, \dots, h_4$  as described in example 9. Each extension  $[0, \epsilon]$  in the set of admissible extensions of *short*,  $\{[0, \epsilon] : \underline{\epsilon} \leq \epsilon \leq \overline{\epsilon}\}$ , identifies one of the valuations  $v_1, \dots, v_5$  given in table 1. Hence, an imprecise definition of *short* based on lower and upper thresholds naturally identifies a set of these valuations i.e. an imprecise valuation. Table 3 shows all the imprecise valuations generated in this manner together with their associated probability. Also shown is the output from applying the strategy  $s$ , as defined in example 9, in each case.

The practical difficulties of determining a joint distribution on  $\mathbb{V} \times \mathbb{S}$ , as discussed in section 4, also apply to joint distributions on  $\mathbb{IV} \times \mathbb{S}$ . Interestingly, there is a parallel solution in which the distribution on  $\mathbb{IV}$  is evaluated by assuming independence between propositional variables. As shown in Aguirre et al. [1] such an independence assumption means that the full distribution on  $\mathbb{IV}$  can be determined from only  $\underline{\mu}(p_i)$  and  $\overline{\mu}(p_i)$  for  $i = 1, \dots, n$ .



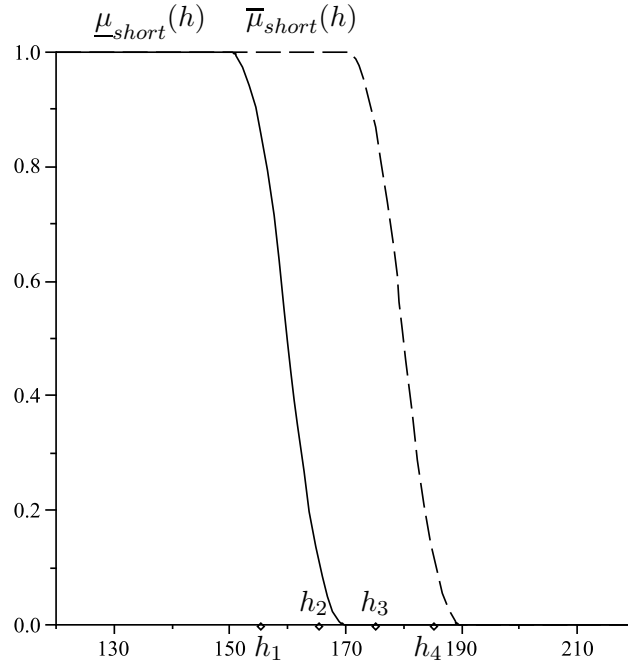


Figure 2: Plot showing the four heights  $h_1, \dots, h_4$  together with the lower and upper membership functions for *short*.

### 6.1 Assertions in the MEL Meta-language

An alternative approach to the problem of assertion given imprecise valuations, is to introduce a meta-language in which agents can explicitly make either strong or weak assertions as appropriate. For example, given an individual with a certain very low height  $h$ , an agent might choose to make a strong assertion describing them as perhaps *definitely short* or *absolutely short*. If, on the other hand,  $h$  is a borderline case of predicate *short* then the agent might instead choose weaker assertions such as *possibly short* or *shortish* etc. In this sub-section we use a simple meta-language called MEL [3]. The MEL language includes two modal operators  $\Diamond$  and  $\Box$  so that for any propositional sentence  $\theta \in \mathcal{SL}$ ,  $\Diamond\theta$  and  $\Box\theta$  denote weak and strong versions of the assertion  $\theta$  respectively. More formally, MEL is defined as follows:

A MEL (Meta-epistemic logic) language [3] for reasoning about incomplete knowledge is defined as a meta-language of  $\mathcal{L}$ , denoted  $\mathcal{ML}$ , as follows. We consider a set of meta-level propositions of the form  $\mathcal{MP} = \{\Box\theta, \Diamond\theta : \theta \in \mathcal{SL}\}$ . The sentences of  $\mathcal{ML}$ , denoted  $\mathcal{SML}$ , are then defined recursively as follows:

- $\mathcal{MP} \subseteq \mathcal{SML}$ .
- If  $\Theta, \Phi \in \mathcal{SML}$  then  $\neg\Theta, \Theta \wedge \Phi, \Theta \vee \Phi \in \mathcal{SML}$ .

**Definition 18.** *Valuations on  $\mathcal{ML}$  (Meta-valuations)*

Given an imprecise valuation  $\Pi \in \mathbb{IV}$  on  $\mathcal{L}$  we define a meta-level valuation  $v^* : S\mathcal{ML} \rightarrow \{0, 1\}$  as follows:

- For  $\theta \in S\mathcal{L}$ ,  $v^*(\Box\theta) = \underline{v}(\theta)$  and  $v^*(\Diamond\theta) = \overline{v}(\theta)$
- For  $\Theta, \Phi \in S\mathcal{ML}$ ,  $v^*(\neg\Theta) = 1 - v^*(\Theta)$ ,  $v^*(\Theta \wedge \Phi) = \min(v^*(\Theta), v^*(\Phi))$  and  $v^*(\Theta \vee \Phi) = \max(v^*(\Theta), v^*(\Phi))$

Let  $\mathbb{MV}$  denote the set of meta-valuations on  $\mathcal{ML}$ .

We now identify a finite set of possible meta-assertions denoted  $A\mathcal{ML} \subseteq S\mathcal{ML}$ .

**Definition 19.** *Meta-level strategies*

A meta-level strategy is a function  $s^* : \mathbb{MV} \rightarrow 2^{A\mathcal{ML}} - \{\emptyset\}$ , such that  $s^*(v) \subseteq A\mathcal{ML}(v^*)$  where  $A\mathcal{ML}(v^*) = \{\Theta \in A\mathcal{ML} : v^*(\Theta) = 1\}$ . Let  $\mathbb{MS}$  denote the set of meta-level strategies.

Now definition 7 can be adapted in a straightforward manner so as to generate a mass function on  $2^{A\mathcal{ML}}$  given a probability distribution on  $\mathbb{MV} \times \mathbb{MS}$ . In addition, given such a mass function on possible meta-level assertions together with a prior distribution on  $A\mathcal{ML}$  then a straightforward adaptation of definition 10 will allow us to determine a posterior distribution on assertions.

**Example 20.** Consider the imprecise definition of short described in example 17, and let  $\mathcal{L}$  have the propositional variables  $\mathcal{P} = \{p_1, \dots, p_4\}$  as given in examples 13 and 17. Let  $\mathcal{ML}$  be the corresponding MEL language and take  $A\mathcal{ML} = \{\Box l, \Diamond l : l \in L\mathcal{L}\}$ . Furthermore, let  $s^*$  be the meta-strategy given by:

$$s^*(v^*) = \{\Diamond p_j, \Box p_r, \Diamond \neg p_k, \Box \neg p_s\} \text{ where } j = \max\{i : v^*(\Diamond p_i) = 1\}, \\ r = \max\{i : v^*(\Box p_i) = 1\}, k = \min\{i : v^*(\Diamond \neg p_i) = 1\} \text{ and } s = \min\{i : v^*(\Box \neg p_i) = 1\}$$

The motivation behind this meta-strategy is to be as informative as possible both with respect to weak and strong assertions. Now given the imprecise valuations  $\Pi_1, \dots, \Pi_9$  shown in table 3 we let  $v_i^*$  be the corresponding meta-valuation defined by  $\Pi_i$  for  $i = 1, \dots, 9$ . Assuming the same semantic uncertainty about the definition of short as in example 17 we take  $w(v_i^*) = w(\Pi_i)$  for  $i = 1, \dots, n$  as given in table 3. We also assume that the agent is certain about the meta-strategy  $s^*$  as defined above so that  $w(v_i^*, s^*) = w(v_i^*)$  for  $i = 1, \dots, n$ . Table 4 shows the meta-valuations  $v_i^* : i = 1, \dots, 9$  together with the associated maximal assertion sets and the value of  $s^*(v_i^*)$ . Hence, given  $w$  on  $\mathbb{MV} \times \mathbb{MS}$

$\Pi$	$w(v^*)$	$AM\mathcal{L}(v^*)$	$s^*(v^*)$
$\Pi_1$	0.015625	$\{\Diamond p_1, \Diamond p_2, \Diamond \neg p_1, \Diamond \neg p_2, \Box \neg p_3, \Diamond \neg p_3, \Box \neg p_4, \Diamond \neg p_4\}$	$\{\Diamond p_2, \Diamond \neg p_1, \Box \neg p_3\}$
$\Pi_2$	0.09375	$\{\Diamond p_1, \Diamond p_2, \Diamond p_3, \Diamond \neg p_1, \Diamond \neg p_2, \Diamond \neg p_3, \Box \neg p_4, \Diamond \neg p_4\}$	$\{\Diamond p_3, \Diamond \neg p_1, \Box \neg p_4\}$
$\Pi_3$	0.015625	$\{\Diamond p_1, \Diamond p_2, \Diamond p_3, \Diamond p_4, \Diamond \neg p_1, \Diamond \neg p_2, \Diamond \neg p_3, \Diamond \neg p_4\}$	$\{\Diamond p_4, \Diamond \neg p_1\}$
$\Pi_4$	0.09375	$\{\Box p_1 \Diamond p_1, \Diamond p_2, \Diamond \neg p_2, \Box \neg p_3, \Diamond \neg p_3, \Box \neg p_4, \Diamond \neg p_4\}$	$\{\Diamond p_2, \Box p_1, \Diamond \neg p_2, \Box \neg p_3\}$
$\Pi_5$	0.5625	$\{\Box p_1, \Diamond p_1, \Diamond p_3, \Diamond \neg p_2, \Diamond \neg p_3, \Box \neg p_4, \Diamond \neg p_4\}$	$\{\Diamond p_3, \Box p_1, \Diamond \neg p_2, \Box \neg p_4\}$
$\Pi_6$	0.09375	$\{\Box p_1, \Diamond p_1, \Diamond p_2, \Diamond p_3, \Diamond p_4, \Diamond \neg p_2, \Diamond \neg p_3, \Diamond \neg p_4\}$	$\{\Diamond p_4, \Box p_1, \Diamond \neg p_2\}$
$\Pi_7$	0.015625	$\{\Box p_1, \Diamond p_1, \Box p_2, \Diamond p_2, \Diamond p_3, \Diamond \neg p_3, \Box \neg p_4, \Diamond \neg p_4\}$	$\{\Diamond p_2, \Box p_2, \Diamond \neg p_3, \Box \neg p_3\}$
$\Pi_8$	0.09375	$\{\Box p_1, \Diamond p_1, \Box p_2, \Diamond p_2, \Diamond p_3, \Diamond \neg p_3, \Box \neg p_4, \Diamond \neg p_4\}$	$\{\Box p_1, \Box p_2, \Box \neg p_4\}$
$\Pi_9$	0.015625	$\{\Box p_1, \Diamond p_1, \Box p_2, \Diamond p_2, \Diamond p_3, \Diamond p_4, \Diamond \neg p_3, \Diamond \neg p_4\}$	$\{\Diamond p_4, \Box p_2, \Diamond \neg p_3\}$

Table 4: Table showing the possible meta-valuations generated by the lower and upper membership functions for the predicate *short*.

as above we generate the following mass function:

$$\begin{aligned}
m := & \{\Diamond p_2, \Diamond \neg p_1, \Box \neg p_3\} : 0.015625, \quad \{\Diamond p_3, \Diamond \neg p_1, \Box \neg p_4\} : 0.09375, \\
& \{\Diamond p_4, \Diamond \neg p_1\} : 0.015625, \quad \{\Diamond p_2, \Box p_1, \Diamond \neg p_2, \Box \neg p_3\} : 0.09375, \\
& \{\Diamond p_3, \Box p_1, \Diamond \neg p_2, \Box \neg p_4\} : 0.5625, \quad \{\Diamond p_4, \Box p_1, \Diamond \neg p_2\} : 0.09375, \\
& \{\Diamond p_2, \Box p_2, \Diamond \neg p_3, \Box \neg p_3\} : 0.015625, \quad \{\Diamond p_3, \Box p_2, \Diamond \neg p_3, \Box \neg p_4\} : 0.09375, \\
& \{\Diamond p_4, \Box p_2, \Diamond \neg p_3\} : 0.015625
\end{aligned}$$

Now suppose that we have a prior probability distribution on  $AM\mathcal{L}$  of the form:

$$\forall l \in L\mathcal{L}, \quad P(\Diamond l) = \frac{\alpha}{2n} \text{ and } P(\Box l) = \frac{1-\alpha}{2n} \text{ for } \alpha \in [0, 1]$$

In this case the resulting posterior distribution generated from the above mass function gives equal maximum probability to  $\Diamond p_3$  and  $\Diamond \neg p_2$  if  $\alpha \geq \frac{1}{2}$  and to  $\Box p_1$  and  $\Box \neg p_4$  if  $\alpha \leq \frac{1}{2}$ .

## 7 Conclusions

In this paper we have proposed a Dempster-Shafer theory based model of assertion for a propositional logic language. Fundamental to our approach is the idea of imprecise assertion strategies which, for a given state of the world, identify a set of possibly optimal assertions. An extended version of the theory then allows for imprecise valuations in addition to imprecise strategies. Two natural extensions of this kind have been proposed; one in which assertions are still taken to be propositional sentences and imprecise strategies are simply extended to sets of valuations, and a second in which assertions have a weak or stronger modifier associated with them and as such correspond to sentences from the MEL meta-language.

As a theme throughout the paper, we have considered how the proposed assertion model can be applied in order to choose between a number of different vague descriptions.

In the case of precise valuations this requires adopting an epistemic approach to vagueness similar to that advocated by Williamson [29]. For imprecise valuations a hybrid approach is required, combining epistemic uncertainty with an underlying supervaluationist truth model. For both approaches we have presented simple examples which are suggestive of how assertion models can be applied to vague propositions, and of what an assertion strategy can look like in this context.

The proposed Dempster-Shafer assertion model is potentially well suited for AI systems which need to generate high-level descriptions of the current state of the world. These include some natural language generation systems and multi-agent communication systems as described in the introduction. More generally, however, an agent's assertions must be seen as forming part of an ongoing dialogue in which context and previous assertions will have a significant impact on the choice of what to say next. In this more general setting an adaptive model is required in which, for example, some treatment of conditioning is given for probabilities defined over a joint space of truth-models and assertion strategies.

## References

- [1] F. Aguirre, C. Jacob, S. Destercke, D. Dubois, M. Sallak, (2013), 'Inclusion/Exclusion Principle for Belief Functions', *Proceedings ISIPTA 2013*, Compiègne, France
- [2] N. Allot, (2006), 'Game Theory and Communication', in (Eds: A. Benz, G. Jäger, R. van Rooij), *Game Theory and Pragmatics*, Palgrave Macmillan, pp123-51.
- [3] M. Banerjee, D. Dubois, (2013), 'A Simple Logic for Reasoning about Incomplete Knowledge', *International Journal of Approximate Reasoning*, Vol. 55, No. 2, pp639-653.
- [4] M. Daniel, (2005), 'Probabilistic Transformation of Belief Functions', *Lecture Notes in Computer Science*, Vol. 3571, pp539-551.
- [5] A. Dempster, (1967), 'Upper and Lower Probability Induced by a Multi-valued Mapping', *The Annals of Mathematical Statistics*, Vol. 38, No. 2, pp325-339.
- [6] D. Dubois, H. Prade, (2001), 'Possibility theory, Probability Theory and Multiple-valued Logics: A Clarification', *Annals of Mathematics and Artificial Intelligence*, Vol. 32, pp35-66.
- [7] H. Eyre, J. Lawry, (2013), 'Communication Success in Vague Category Games', *Proceedings of the AISB 2013 convention*.
- [8] H. Eyre, J. Lawry, (2014), 'Language Games with Vague Categories and Negations', *Adaptive Behaviour*, Vol. 22, No. 5, pp289-303

- [9] K. Fine, (1975), ‘Vagueness, Truth and Logic’, *Synthese*, Vol. 30, pp265-300.
- [10] E. Goldberg, N. Driedger, R.I. Kittredge, (1994), ‘Using Natural-Language Processing to Produce Weather Forecasts’, *IEEE Expert: Intelligent Systems and their Applications*, Vol. 9, No. 2, pp45-53.
- [11] H.P. Grice, (1975), ‘Logic and Conversation’, in (Eds: P. Cole, J. Morgan), *Syntax and Semantics*, Vol. 3, Academic Press, pp22-40.
- [12] P. Hansen, B. Jaumard, M. Poggi de Aragão, F. Chauny, S. Perron, (2000), ‘Probabilistic Satisfiability with Imprecise Probabilities’, *International Journal of Approximate Reasoning*, Vol. 24, pp171-189.
- [13] J. Lawry, (2008), ‘An Overview of Computing with Words using Label Semantics’, in (Eds: H. Bustince, F. Herrera, J. Montero), *Fuzzy Sets and their Extensions: Representation, Aggregation and Models*, Springer, pp65-87.
- [14] J. Lawry, (2008), ‘Appropriateness Measures: An Uncertainty Model for Vague Concepts’, *Synthese*, Vol. 161, No. 2, pp255-269.
- [15] J. Lawry, Y. Tang, (2009), ‘Uncertainty Modelling for Vague Concepts: A Prototype Theory Approach’, *Artificial Intelligence*, Vol. 173, pp1539-1558.
- [16] J. Lawry, Y. Tang, (2012), ‘On Truth-gaps, Bipolar Belief and the Assertability of Vague Propositions’, *Artificial Intelligence*, Vol. 191-192, pp20-41.
- [17] I. Levi, (1980), *The Enterprise of Knowledge*, MIT press
- [18] D. Lewis, (1969), *Convention: A Philosophical Study*, Harvard University Press
- [19] H.T. Nguyen, (1978), ‘On Random Sets and Belief Functions’, *Journal of Mathematical Analysis and Applications*, Vol. 65, pp531-542
- [20] P. Parikh, (2000), ‘Communication, Meaning, and Interpretation’, *Linguistics and Philosophy*, Vol. 23, pp185-212
- [21] R. Parikh, (1996), ‘Vague predicates and language games’, *Theoria (Spain)*, Vol. 11, No. 27 pp97-107.
- [22] J.B. Paris, (1994), *The Uncertain Reasoners Companion: A Mathematical Perspective*, Cambridge University Press
- [23] F. Portet, E. Reiter, A. Gatt, J. Hunter, S. Sripada, Y. Freer, C. Sykes, (2009), ‘Automatic Generation of Textual Summaries from Neonatal Intensive Care Data’, *Artificial Intelligence*, Vol. 173, No. 7-8, pp789-816.

- [24] E. Reiter, R. Dale, (2006), *Building Natural Language Generation Systems*, Cambridge University Press.
- [25] G. Shafer, (1976), *A Mathematical Theory of Evidence*, Princeton University Press.
- [26] P Smets, (1990), ‘Constructing the Pignistic Probability Function in a Context of Uncertainty’, in *Uncertainty in Artificial Intelligence 5*, (Ed. M. Henrion), North Holland, Amsterdam, pp29-39
- [27] L. Steels, (2001), ‘Language Games for Autonomous Robots’, *IEEE Intelligent Systems*, Vol.16, No. 5, pp16-22.
- [28] K. van Deemter, (2009), ‘Utility and Language Generation: The Case of Vagueness’, *Journal of Philosophical Logic*, Vol. 38, pp607-632.
- [29] T. Williamson, (1994), *Vagueness*, Routledge.